

METHOD FOR DETECTING SHOT AND DEVICE FOR RECORDING/ REPRODUCING REPRESENTATIVE PICTURE

Publication number: JP7236115

Publication date: 1995-09-05

Inventor: GOSHIMA YUKIE; AKAHORI HIROSHI; FUJIMOTO MAKOTO

Applicant: MATSUSHITA ELECTRIC IND CO LTD

Classification:

- International: H04N5/7826; G06T7/20; H04N5/765; H04N5/7824; G06T7/20; H04N5/765; (IPC-1-7): H04N5/7826; G06T7/20; H04N5/765

- European:

Application number: JP19940026507 19940224

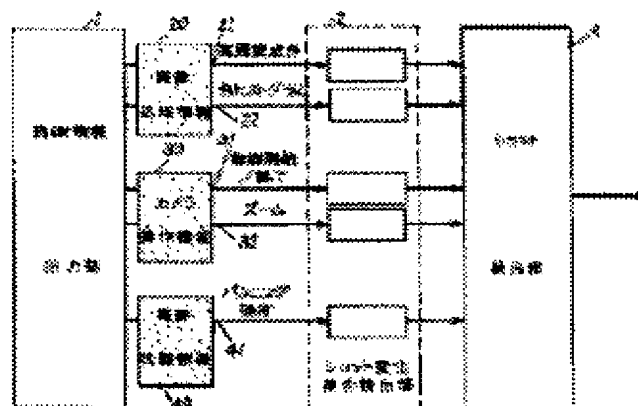
Priority number(s): JP19940026507 19940224

Report a data error here

Abstract of JP7236115

PURPOSE: To collect moving pictures in units corresponding to the contents of pictures.

CONSTITUTION: A picture information output part 1 outputs the camera operation information of a camera with which a user operates for photographing, picture processing information obtained in processing a picked-up image or photographing state information obtained in processing a signal inputted from a sensor. A shot variation degree detecting part 2 detects the degree of variation in a picture in each input information by previously setting up changing rules specified by various information outputted from the output part 1 and comparing successively inputted information with the rules. A shot detecting part 3 detects a moving picture obtained by continuously photographing a specific subject by a photographer or a moving image continuously photographed by the photographer under a specific viewing angle, photographing condition or photographing method as one shot based upon an output from the detecting part 2 and outputs the shot section.



Data supplied from the esp@cenet database - Worldwide

Method for Detecting Shot and Device for Recording / Reproducing
Representative Picture

[0022] Here, image processing information is collective denotation of information extracted and processed automatically or through involvement of a person based on a signal of a video image picked up by an image pick-up device. Fig. 3 shows, by way of example, high-frequency component magnitude 21 of the video image signal obtained for focus control, color histogram 22 in the screen, and motion vector 23 of each part of the screen. Such information is obtained based on input of video image signal 5 from camera 4, by high-frequency component detection unit 24, color histogram detection unit 25, and motion vector detection unit 26. In addition, the image processing information also includes an inter-frame differential value representing a difference between frames in a luminance signal or a color signal, information on a position or a size of a subject region from a video image signal, or the like.

[0036] In the embodiment above, a method of outputting information directly from a camera during image pick-up has been described as a specific example of image information output unit 1, with reference to Fig. 3. On the other hand, similar processing can be performed with another method, that is, by recording information from a camera in a recording medium during image pick-up and reading the information from the recording medium during shot detection. This method will be described with reference to Fig. 7. Fig. 7 shows camera 4, video image signal 5, recording medium 6, encode processing unit 7, decode processing unit 8, image processing unit 9, image processing information 20, camera operation information 30, and image pick-up state information 40, with image information output unit 1 being defined by a frame shown with a dashed line.

[0037] An operation of each unit will be described as follows. Recording medium 6 includes a switch switched between "during recording" and "during reproduction". Here, "during recording" refers to a period of time during which various types of information are recorded in a recording medium while image pick-up by camera 4 is being carried out, and "during reproduction" refers to a period of time during which various types of information recorded in the recording medium are output. Initially, during image pick-up by the user, the switch of recording medium 6 is set to a "during recording" mode. Video image signal 5 output from camera 4 during image pick-up, image processing information, camera operation information, and image pick-up state information are subjected to encode processing or format-matching processing in encode processing unit 7 and thereafter stored in recording medium 6. Then, during shot detection, the switch of recording medium 6 is set to a "during reproduction" mode. The information stored in recording medium 6 is read by decode processing unit 8 as respective types of information and sent to shot variation detection unit 2 as outputs from image information output unit 1. Here, the same result is obtained even when video image signal 5 output from decode processing unit 8 is processed by image processing unit 9 and the resultant signal is output together with other image processing information.

[0038] Thus, as in the case of directly obtaining the information from the camera, the various types of information can be output also by once storing the various types of information in the recording medium.

Fig. 3 is a diagram showing a configuration of an image information output unit in the first embodiment of the shot detection method according to the present invention.

Fig. 7 is a diagram showing a configuration different from that in Fig. 3, in connection with the image information output unit in the first embodiment of the shot detection method according to the present invention.

Fig. 3

4	camera
5	video image signal
24	high-frequency component detection unit
25	color histogram detection unit
26	motion vector detection unit
ボタン押下信号	signal indicating that the button has been pressed
31	recording start/end
32	zoom
41	angular velocity sensor output (panning speed)
42	zoom focal length
43	aperture sensor output
44	focus length
1	image information output unit
20	image processing information
30	camera operation information
40	image pick-up state information

Fig. 7

6	recording medium
録画時	during recording
再生時	during reproduction
7	encode processing unit
8	decode processing unit
5	video image signal
画像処理情報	image processing information
カメラ操作情報	camera operation information

撮影状態情報 image pick-up state information

4 camera

映像信号 video image signal

9 image processing unit

画像処理情報 image processing information

カメラ操作情報 camera operation information

撮影状態情報 image pick-up state information

20 image processing information

30 camera operation information

40 image pick-up state information

1 image information output unit

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-236115

(43) 公開日 平成7年(1995)9月5日

(51) Int.Cl. ⁹	識別記号	序内整理番号	F I	技術表示箇所
H 0 4 N	5/7826			
G 0 6 T	7/20			
H 0 4 N	5/765			
		9061-5L	H 0 4 N 5/ 782 G 0 6 F 15/ 70	A 4 0 5
審査請求 未請求 請求項の数 6 O L (全 14 頁) 最終頁に続く				

(21) 出願番号 特願平6-26507

(22) 出願日 平成6年(1994)2月24日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 五島 雪絵

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 赤堀 裕志

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 藤本 眞

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

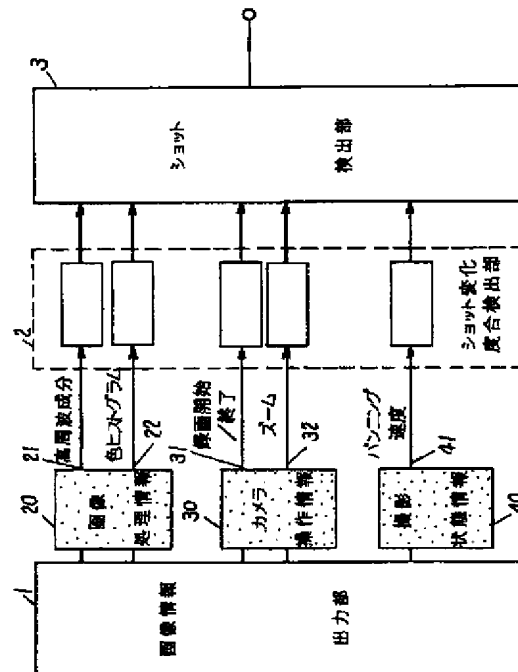
(74) 代理人 弁理士 小銀治 明 (外2名)

(54) 【発明の名称】 ショット検出方法および代表画像記録・表示装置

(57) 【要約】

【目的】 画像内容に応じた単位で動画像をまとめること。

【構成】 画像情報出力部1は、ユーザが撮影時にカメラを操作したカメラ操作情報や、撮像した画像を処理して得られた画像処理情報や、センサーからの信号を処理して得られた撮影中の撮影状態情報を出力する。ショット変化度合検出部2は、画像内容が移行するときに画像情報出力部1から出力される各種情報が示す変化規則を予め設定しておき、順次入力される各種情報を変化規則と比較することで、入力情報毎に画面の変動する度合を検出する。ショット検出部3では、ショット変化度合検出部2からの出力を基に、撮影者が特定の被写体を撮影しつづけた動画像、及び撮影者が特定の画角や撮影条件や撮影方法で撮影し続けた動画像を1ショットとして検出し、そのショットの区間を出力する。



1

【特許請求の範囲】

【請求項 1】 撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中で、撮影者が特定の被写体を撮影しつづけた動画像、及び撮影者が特定の画角や撮影条件や撮影方法で撮影し続けた動画像をショットとし、撮像装置のズームや撮影開始操作などのカメラ操作情報と撮像した画像を処理して得られた画像処理情報と撮影中のセンサーからの信号を処理して得られた撮影状態情報の少なくとも一つを入力情報とし、動画像が一方のショットから別のショットへ移行する際の前記入力情報それぞれの変化の規則を予め選出しておき、前記入力情報それぞれに関して前記変化の規則と合っている度合をショット変化度合として検出し、少なくとも 1 つの前記ショット変化度合をもとにして、動画像中のショットを検出することを特徴とするショット検出方法。

【請求項 2】 前記カメラ操作情報および前記画像処理情報および前記撮影状態情報の前記入力情報は、カメラでの撮影の際、または予め前記入力情報の蓄積された記録媒体を再生する際、カメラまたは記録媒体から前記入力情報を直接読み出す、またはカメラまたは記録媒体から出力される映像信号を基に前記入力情報を推定することによって獲得することを特徴とする請求項 1 記載のショット検出方法。

【請求項 3】 動画像が一方のショットから別のショットへ移行する際に前記入力情報が大きく変動することを利用し、各前記入力情報に関して隣接フレーム間での相関値または時間軸の微分フィルタの出力をもとに前記ショット変化度合を検出することを特徴とする請求項 1 記載のショット検出方法。

【請求項 4】 入力情報に関して、隣接フレーム間での相関値または時間軸の微分フィルタの出力、または前記入力情報の大きさ、または前記入力情報が微小な変動のまま継続している時間をもとに前記ショット変化度合を検出することを特徴とする請求項 1 記載のショット検出方法。

【請求項 5】 撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中で、撮影者が特定の被写体を撮影しつづけた動画像、及び撮影者が特定の画角や撮影条件や撮影方法で撮影し続けた動画像をショットとし、動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を取り込むカメラ操作情報獲得手段と、撮像した画像を処理して得られた画像処理情報を取り込む画像処理情報獲得手段と、センサーからの信号を処理して得られた撮影中の撮影状態情報を取り込む撮影状態情報獲得手段の 3 手段のうち少なくとも 1 つを備えて前記カメラ操作情報または前記画像処理情報または前記撮影状態情報を出力する画像情報出力部と、前記画像情報出力部からの複数の出力それぞれに関して、予め設定された変化の規則に合っている度合をショット

2

変化度合として検出するショット変化度合検出部と、前記ショット変化度合検出部からの出力の少なくとも 1 つを基にして、動画像中のショットを検出するショット検出部と、前記ショット検出部で求められたショットに属する画像から代表的な画像を抽出する代表画像抽出部と、動画像の映像信号を取り込む映像情報獲得部と、前記代表画像抽出部で抽出した代表画像の映像信号を前記映像情報獲得部から入力し、画像記録媒体に記録、または画像表示装置に表示する画像記録・表示部を備えたことを特徴とする代表画像記録・表示装置。

【請求項 6】 代表画像抽出部は、前記ショット検出部で検出された各ショットに対して、同一ショットに属する画像の各種情報を入力して各ショットの代表情報を求めるショット内代表情報検出部と、前記ショット内代表情報検出部で求められたショットの代表情報を蓄えるメモリと、ショット間で代表情報を比較して、比較結果を基にショットを選び、選ばれたショットに属する画像から代表画像を抽出する代表情報比較部から構成されることを特徴とする請求項 5 記載の代表画像記録・表示装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、動画像の早見、検索、編集などのために、画像内容に応じた単位で動画像をまとめる方法、及び、まとめられた動画像を基に代表的な画像を自動抽出し、その画像を記録・表示する装置に関するものである。

【0002】

【従来の技術】 従来、動画像の区切りを検出する技術や、動画像の区切りを編集作業に利用する技術が多数提案されてきた。

【0003】 動画像の区切りを映像信号から自動検出する方法としては、例えば、特開平 3-214364 号公報に開示されたシーン変化検出手法がある。これは、隣接フレーム間でヒストグラムを比較することにより、シーンが変化したかどうか判定するものである。ここでのシーンの変化は、撮影・録画の開始/終了時点や、編集作業によって別々のシーンを連結した時のつなぎ目に対応している。また、V I S S (VHS Index Search System) では、録画開始時に、ビデオテープに V I S S 信号を自動記録し、高速頭出しなどのタグとして利用している。

【0004】 一方、録画の開始点だけでなく、好みの画像の時点にタグ付けする、といった柔軟な技術も必要とされている。例えば、前述の V I S S では、ユーザが見たい場面に対しても V I S S 信号を記録できるので、ビデオテープ上に記録された V I S S 信号を利用して、イントロサーチと呼ばれる早送り再生を行うことができる。イントロサーチとは、早送り中に V I S S 信号を見つけると、ある時間だけ再生状態にし、その後再び早送りするという動作をテープの終わりまで繰り返すもので

ある。

【0005】さらに、「好みの画像を選択してVISS信号を付加する」というユーザの手間を省くために、代表的な画像を自動抽出する方法も提案されている。例えば、特開平5-147337では、1カット（録画開始操作をしてから録画終了操作をするまでの間に連続して撮影された動画像）の中から、代表的な画像を自動的に抽出する静止画像自動抽出方法が開示されている。ここで、代表的な画像は、撮影者の意図、撮影された画像の状態、被写体の状態をもとに評価して選び出され、動画像の早見や検索などに利用される。

【0006】

【発明が解決しようとする課題】しかしながら、上記の静止画像自動抽出方法も、録画開始から終了までの動画像ごとに代表画像を抽出するので、基本的に代表画像の数は、動画像全体に含まれる録画開始／終了の区切りの数に依存する。

【0007】例えば、映画のようにカットの多い動画像の場合には、代表画像が多くなる。この代表画像は少数の画像の中から選ばれるので、各カットの画像内容をよく表すような画像が選ばれ易い。一方、ビデオカメラで長時間、撮影しつづけた映像の場合には、どんなに画像内容が変化しても、録画開始／終了の区切りがないので、ほとんど代表画像を抽出できない。従って、代表画像だけ見ても、動画像全体の内容がわからない場合もある。

【0008】ここで具体例として、撮影開始から終了までに、2種類の被写体を撮影するような場合を考えてみる。図2の画像A～画像Pは、長時間撮影した動画像から、一部分をぬきだしたものである。撮影者はまず、「黄色の車」を撮影して（画像A、B）いったん撮影を止め、撮影開始後、「赤い服の人物」を撮影し（画像C～画像F）、そのままカメラをパンニングして（画像G～画像I）、「茶色の屋根の小屋」を撮影（画像J～画像N）、ここで撮影を中止し、再開後は「高層ビル」を撮影している（画像O、P）。録画開始／終了の区切りは、画像Bと画像Cの間、画像Nと画像Oの間であり、画像C～Nが一まとまりの画像とみなされる。しかし、画像内容から考えると画像C～Nの中で、「赤い服の人物」の部分と「茶色の屋根の小屋」の部分は別の場面であり、それぞれで代表画像を選んだ方がよい。

【0009】このように、撮影や編集操作による区切りで動画像をまとめても、まとまりによって画像内容の変化する度合が異なるため、内容を表現するのに必要な代表画像の数が増減し、従来方法では対応できないことがわかる。従って、画像内容に応じた単位で動画像をまとめて取り扱うことが必要になる。

【0010】本発明はかかる点に鑑み、画像内容に応じて動画像をまとめる手法、およびまとめられた動画像を基に代表的画像を自動抽出し、記録表示する装置を提供

することを目的とする。

【0011】

【課題を解決するための手段】上記課題を解決するために、本発明のショット検出方法は、撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中で、撮影者が特定の被写体を撮影しつづけた動画像、及び撮影者が特定の面角や撮影条件や撮影方法で撮影し続けた動画像をショットとし、動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を取り込むカメラ操作情報獲得手段と、撮像した画像を処理して得られた画像処理情報を取り込む画像処理情報獲得手段と、センサーからの信号を処理して得られた撮影中の撮影状態情報を取り込む撮影状態情報獲得手段の3つの手段のうち少なくとも1つを備えて前記カメラ操作情報または前記画像処理情報または前記撮影状態情報を出力する画像情報出力部と、前記画像情報出力部からの複数の出力それぞれに関して、予め設定された変化の規則に合っている度合をショット変化度合として検出するショット変化度合検出部と、前記ショット変化度合検出部からの出力の少なくとも1つを基にして、動画像中のショットを検出するショット検出部を備える。

【0012】さらに本発明のショット内代表画像記録・表示装置は、前記ショット検出部で求められたショットに属する画像から代表的な画像を抽出する代表画像抽出部と、動画像の映像信号を取り込む映像情報獲得部と、前記代表画像抽出部で抽出した代表画像の映像信号を前記映像情報獲得部から入力し、画像記録媒体に記録、または画像表示装置に表示する画像記録・表示部を備える。

【0013】

【作用】以上のような構成において、カメラ操作情報や画像処理情報や撮影状態情報を入力情報とし、動画像の内容が変化する場合の入力情報それぞれの変化規則を予め選出しておき、変化規則と入力情報を比較することで、撮影者が特定の被写体を撮影しつづけた動画像や特定の面角、撮影条件で撮影しつづけた動画像をショットとして検出できる。これによって、録画開始から終了までの動画像の単位にとらわれず、動画像を画像内容に応じた任意個数のかたまりにまとめることができる。

【0014】また、上記方法で検出されたショットを基に代表的画像を選ぶことにより、動画像の内容の変化度合に応じて必要個数の代表画像が得られるようになり、さらに、抽出した代表画像の情報を記録・表示することで、できるだけ少ない画像で動画像全体の内容をよく表せるようになる。

【0015】

【実施例】

（実施例1）まず、本発明のショット検出方法の第1の実施例について説明する。ここで、「ショット」とは、1カットの画像を撮影している間に、パンニングやズー

ミングなどのカメラ操作を続けたり、特定の被写体を撮りつづけていたりして、撮影者が1つの場面として意図的に連続撮影した動画像のかたまりとする。それに対して、「カット」は、カメラにおいて録画開始操作をしてから録画終了操作をするまでの間に連続して撮影された動画像のかたまりとする。

【0016】例えば、図2の動画像に対して、カットの単位で区切ると、画像Bまでが「黄色い車」の写ったカット、画像C～Nが「赤い服の人物」と「茶色い屋根の小屋」の両方の写ったカット、画像Oからは「高層ビル」の写ったカットが始まっている。一方、ショットについて考えると、画像C～Nのカットの中に、

- ・「赤い服の人物」をアップショットで撮影した区間（画像E～F付近）
 - ・「茶色い屋根の小屋」をややロングショットで撮影した区間（画像M～N付近）
- の2つのショットが含まれることがわかる。

【0017】まず、図1を基に本発明のショット検出方法の概要を説明する。図に示すように、本発明のショット検出方法は、画像情報出力部1、ショット変化度合検出部2、ショット検出部3の3つの部分で実現されている。

【0018】画像情報出力部1は、映像信号を処理した情報、撮影時にユーザが行ったカメラ操作の情報、センサの出力情報を出力する。この3種類の情報に関しては、以後、それぞれの種類の情報を総称して、画像処理情報、カメラ操作情報、撮影状態情報と呼ぶことにする。3情報の詳しい説明は後述する。図1では、画像処理情報20の例として、高周波成分（の大きさ）21と色ヒストグラム22を、カメラ操作情報30の例として、録画開始／終了31とズーム（倍率）32を、撮影状態情報40の例として、パンニング速度41を挙げている。

【0019】ショット変化度合検出部2は、画像情報出力部1からの情報を入力として、画像内容の変化する度合を検出する。ショット検出部3では、ショット変化度合検出部2からの結果をもとに、動画像中でショットと判定される区間を検出する。

【0020】続いて、画像情報出力部1、ショット変化度合検出部2、ショット検出部3の各部の動作について詳細に説明する。

【0021】始めに、図3を用いて画像情報出力部1について説明する。図において、4はカメラ、5は映像信号、20は画像処理情報、21は高周波成分の大きさ、22は色ヒストグラム、23は動きベクトル、24は高周波成分検出部、25は色ヒストグラム検出部、26は動きベクトル検出部、30はカメラ操作情報、31は録画開始／終了のボタン押下信号、32はズーム倍率、33はオート／マニュアルのモードのボタン設定、40は撮影状態情報、41は角速度センサの出力（パンニング

時の速度）、42はズーム倍率を換算するためのレンズ焦点距離、43は絞り開度センサの出力、44はフォーカス距離である。図3に示すように、画像情報出力部1は、カメラと多数の出力端子からなっており、ビデオカメラで撮影中、画像処理情報20、カメラ操作情報30、撮影状態情報40を順次出力する。

【0022】ここで画像処理情報は、撮像素子で撮像した映像信号をもとにして自動的にもしくは人間が関与して抽出処理した情報の総称である。図3では例として、フォーカス制御を行うために求めた映像信号の高周波成分の大きさ21、画面内の色ヒストグラム22、画面各所の動きベクトル23を示している。この情報はすべて、カメラ4から映像信号5を入力し、それぞれ高周波成分検出部24、色ヒストグラム検出部25、動きベクトル検出部26において求められる。その他、フレーム間における輝度信号や色信号の差異を求めたフレーム間差分値、あるいは映像信号から被写体領域の位置や大きさの情報なども画像処理情報に含まれる。

【0023】また、カメラ操作情報は、ユーザがビデオカメラで撮影する際に行った、ボタン操作をもとにした情報の総称である。図3では、例として、撮影の開始／終了のボタン操作により撮影の開始点／終了点を表す情報31、撮影中のズーム操作から換算されたズーム倍率を表す情報32、Auto/Manualの設定の情報33を示している。

【0024】また撮影状態情報は、カメラの撮影中の状態をセンサーなどで検出した情報の総称である。図3では、例として、角速度センサーによって検出したパンニング速度の情報41、撮影時のズーム倍率を示すレンズ焦点距離42、絞り開度センサによって検出した絞り開度情報43、フォーカス距離44を示している。

【0025】次に、ショット変化度合検出部2とショット検出部3について説明する。ただしここでは、例として、高周波成分（の大きさ）21、色ヒストグラム22、録画開始／終了のボタン押下信号31、ズーム倍率32、パンニング速度41の5つの情報が入力する場合に限定して説明する。

【0026】まず、図2の動画像を例にとり、ショット変化度合検出部2、ショット検出部3において、ショットを検出する原理を説明する。図5は、図2の各画像に対して、録画開始／終了、ズーム倍率、画面内の高周波成分の大きさ、色ヒストグラム、パンニングの速度、の5つの情報をプロットしたものである。ただし、色ヒストグラムは多数の色の中から茶色、赤色、黄緑に対する頻度だけを抽出して表示した。また、パンニング速度は、水平方向の速度に限定している。図からわかるように、ショットの区間（E～F、M～N）は、全ての情報が一定値で安定しているのに対し、1つのショットから別のショットへ移る区間（C～E、F～M）では少なくとも1つの情報が変化していることがわかる。このこ

とから、入力情報の変化度合を調べれば、現在処理中の画像が、ショットの区間中なのか、それとも別のショットに移行中なのか、わかるはずである。ここで、「処理中の画像が別のショットに移行中である」という信頼度をショット変化度合と呼ぶことにし、以下、入力される情報の変化度合を、各情報のショット変化度合として検出し、各情報に関するショット変化度合からショットの区間を特定する方法を説明する。

【0027】図4は、ショット変化度合検出部2、ショット検出部3の具体的な構成図である。図において、51は微分フィルタ、52は絶対値処理部、53はローパスフィルタ、54はゲイン調整部、56は最大値検出部、57は状態判定部、58はショット区間出力部、59はカウンタである。

【0028】まず、ショット変化度合検出部2では、入力される情報それぞれのショット変化度合を検出する。ここで、ショット変化度合は $[0, \max]$ (\max : 定数)の範囲の値に正規化されており、値が大きいほど、処理中の画像が別のショットに移行中の可能性が高いことを示すものとする。

【0029】ショット変化度合の検出方法は、入力情報によって多少異なっている。高周波成分(の大きさ)21、ズーム倍率32、パンニング速度41の情報に関しては、まず微分フィルタ51を通し、絶対値処理部52で絶対値を求めることにより、各情報の時間的変動の大きさを求める。さらに、細かいノイズの影響を少なく *

$M > TH$ のとき、現在処理中の画像はショット外である

(別のショットへ移行中である)

… (A)

$M \leq TH$ のとき、現在処理中の画像はショット区間内にある … (B)

ショット区間出力部58では、状態判定部57の出力(A or B)に応じて以下のような処理を行う。

・状態判定部57の出力がAのとき、カウンタ59の内容をCを読み込む。

【0032】 $C > 0$ の場合、Cを出力し、その後カウンタ59の内容を0にリセットする。

・状態判定部57の出力がBのとき、カウンタ59の内容をインクリメントする。

【0033】ただし、初期状態で、カウンタは0に設定されているとする。ショット終了時点で出力される値Cは、ショット継続中に処理した回数を表している。従って、ショット検出部3から値Cが出力された時、その出力を基に、直前に終了したショットの区間を特定することができる。

【0034】ここで、以上説明した構成が、実際の画像でどのように動作するか調べるため、図2の動画像を例にとって情報の流れを追ってみる。図6は、ショット変化度合検出部2とショット検出部3の内部のブロックの出力の時間変化を示した図である。図の上部のグラフはショット変化度合検出部2に対応しており、その中で実線は微分フィルタ51またはヒストグラム相関検出部5

*し、大域的な変化だけを検出するため、ローパスフィルタ53を通す。最後にゲイン調整部54において、各情報固有のゲインをかけて正規化し、出力がショット変化度合の範囲内の値をとるように調整する。また色ヒストグラムのように、多次元のパラメータに関しては、隣接フレーム間での相関計算を行う。色ヒストグラム22の場合、ヒストグラム相関検出部55において、ヒストグラム差分などの相関計算を行う。さらに他の情報と同様に、ローパスフィルタ53を通し、ゲイン調整部54において正規化する。ただし、ここでの正規化は、相関が低いほど大きい値を出力するように調整する。また、録画開始/終了のボタン押下信号31に関しては、ゲイン調整部54において、カットの切れ目のとき \max 、切れ目でないとき0の値を出力するよう、調整する。以上のようにして、ショット変化度合検出部2では、入力情報毎にショット変化度合を検出する。

【0030】一方、ショット検出部3では、各情報に関するショット変化度合をもとに、ショットの区間を検出する。まず、最大値検出部56では、ショット変化度合検出部2から同時刻に出力されるすべてのショット変化度合を比較し、その最大値Mを求める。状態判定部57では、求めた最大値Mを所定のしきい値THと比較し、比較結果をもとに、現在処理中の画像の状態を以下のように判定する。

【0031】

5の出力を表し、破線はショット変化度合検出部2の最終的な出力(ショット変化度合)を表している。また、図6の下部のショット検出部3に対応するグラフでは、ショット変化度合検出部2の出力(破線)とその最大値(太い実線)、しきい値を重ねて示した。また、取り扱う入力情報は、録画開始/終了31、ズーム倍率32、画面内の高周波成分の大きさ21、色ヒストグラム22、パンニングの速度41、の5つに限定し、さらに色ヒストグラムは3色のみ、パンニングの速度は水平方向に限定して表示してある。図6において、各入力情報に対するショット変化度合は、ショット以外の区間(C~E、F~M)で大きな値を出力している。また、ショット変化度合の最大値は、ショット以外の区間でしきい値を越えていることがわかる。従って、FとNの画像を処理する際、ショット検出部3からそれぞれE~F、M~Nの区間中の処理回数(L1、L2)が出力され、L1、L2により、E~F、M~Nがショットの区間であると特定できる。

【0035】以上のように、カメラ操作情報または画像処理情報または撮影状態情報を入力とし、各情報に関するショット変化度合を検出することにより、ショットの

区間が特定でき、動画を画像内容に応じた単位でまとめることができる。

【0036】なお、上記実施例では、画像情報出力部1の具体例としては、図3を用いて、撮影中、カメラから直接情報を出力する方法を説明した。しかし、別の方法として、撮影中はカメラからの情報を記録媒体に記録し、ショット検出を行う際に、記録媒体から情報を読み出しても同様の処理が行える。この方法を図7を用いて説明する。図において、4はカメラ、5は映像信号、6は記録媒体、7は符号化処理部、8は復号化処理部、9は画像処理部、20は画像処理情報、30はカメラ操作情報、40は撮影状態情報であり、破線枠内が画像情報出力部1に対応している。

【0037】以下に各部の動作を説明する。記録媒体6は、録画時と再生時で切り替わるスイッチを持っている。ここで、「録画時」とは、カメラ4で撮影しながら、各種情報を記録媒体に記録する期間であり、「再生時」とは、記録媒体に記録された各種情報を出力する期間である。まず、ユーザが撮影する際、記録媒体6のスイッチは「録画時」のモードに設定されている。撮影中にカメラ4から出力された映像信号5、画像処理情報、カメラ操作情報、撮影状態情報は、符号化処理部7において、符号化の処理やフォーマット合わせの処理が施され、記録媒体6に蓄積される。その後、ショットを検出する際に、記録媒体6のスイッチは「再生時」のモードに設定される。記録媒体6に蓄えられた情報は、復号化処理部8でそれぞれの情報として読み出され、画像情報出力部1の出力として、ショット変化度合検出部2に送られる。ここで、復号化処理部8をから出力された映像信号5を画像処理部9で加工し、その結果を他の画像処理情報と合わせて出力しても同様である。

【0038】このように、各種情報をいったん記録媒体に蓄積しても、カメラから直接得ると同様にして、各種情報を出力できる。

【0039】以上の実施例では、画像処理情報、カメラ操作情報、撮影状態情報を、カメラから直接出力したり、いったん記録媒体に蓄積して後から読みだしたりする場合について説明したが、いずれにしても、必要なすべての情報は、もともとカメラから与えられていた。しかし、これら情報の一部、もしくは全部が与えられない場合でも、カメラあるいは記録媒体から出力される映像信号を処理することによって、欠如した情報に相当する情報を獲得し、獲得した情報をもとにショットを検出することができる。これについて、以下の第2の実施例のショット検出方法で詳細に説明する。

【0040】(実施例2) 第2の実施例は、映像信号のみから画像処理情報、カメラ操作情報、撮影状態情報を獲得し、獲得した情報からショットを検出するものである。本実施例の全体構成は、第1の実施例で用いた図1と同じであるが、画像情報出力部1の具体的な構成が異

なっている。以下、図8を用いて、本実施例の画像情報出力部1において映像信号から各種情報を検出・出力する方法を説明する。ただし、ここで説明するのは、画像処理情報に関して高周波成分と色ヒストグラム、カメラ操作情報に関しては録画開始/終了の情報とズーム倍率、撮影状態情報に関してはパンニング速度、の5つの情報に限定している。

【0041】図8で、5は映像信号、26は高周波成分検出部、21は高周波成分(の大きさ)の情報、27は色ヒストグラム検出部、22は色ヒストグラムの情報、60はフレーム間差分値検出部、61はメモリ、62は変化量検出部、63はカットチェンジ検出部、31は録画開始/終了の情報、64はカメラワーク検出部、65は動きベクトル検出部、66はカメラワークパラメータ推定部、32はズーム倍率の情報、41はパンニング速度の情報である。以上の構成における各部の動作について以下に説明する。

【0042】まず、画像処理情報に関しては、高周波成分検出部24、色ヒストグラム検出部25において、それぞれ高周波成分21、色ヒストグラム22の情報を検出する。ここで、高周波成分検出部24と色ヒストグラム検出部25は、図3の各部と同じであり、既に第1の実施例で述べたので、説明は省略する。

【0043】次に、カメラ操作情報の内、録画開始/終了31の情報を検出する方法を説明する。録画開始/終了の情報の検出は、フレーム間差分値検出部60とカットチェンジ検出部63で行われ、さらにフレーム間差分値検出部60は、映像信号を1フレーム遅延させるためのメモリ61と、連続するフレーム間で映像信号の差分を求める変化量検出部62から構成されている。連続するフレーム間での差を求める信号は、輝度値やRGB値等を用い、変化量検出部62において画素単位で連続するフレーム間の映像信号の差分演算を行い、画素ごとの差分値の総和を求めてフレーム間差分値として出力する。カットチェンジ検出部63は、フレーム間差分値検出部60で求めたフレーム間差分値に対してしきい値処理をする。すなわち、所定のしきい値とフレーム間差分値との比較を行い、フレーム間差分値がしきい値より大きい場合は、2枚のフレーム間で画像内容が大きく変化していると考えて、その部分でカットチェンジがあったと判断する。ここで検出したカットチェンジは、撮影の開始や終了時点に対応するので、カットチェンジの有無の信号は、カメラのボタン押下信号として得られる録画開始/終了の情報に相当する。

【0044】次に、カメラ操作情報の中のズーム倍率32の情報と、撮影状態情報のパンニング速度41の情報を検出する方法を説明する。この2つの情報は、カメラワーク検出部64において検出され、さらにカメラワーク検出部64は、動きベクトル検出部65とカメラワークパラメータ推定部66に分かれる。まず、動きベクトル

ル検出部65は、画面内で複数の座標位置を設定し、隣接フレームとの画素値の比較により、各座標位置での動きベクトルを検出する。カメラワークパラメータ推定部66では、検出された動きベクトルを基に、カメラの水平、垂直方向の変化（パンニング、チルティング）や、カメラ画角の変化（ズームング）、カメラの水平・垂直・前後の位置の変化（トラッキング、プーミング、ドリーイング）等のカメラワークを推定する。ズーム倍率32とパンニング速度41は、カメラワークの1つとして推定されるので、カメラ操作情報や撮影状態情報に相当する情報が得られる。ここで、動きベクトル検出部65、カメラワークパラメータ推定部66の詳細な動作は、例えば特開平4-317267で説明されているので、ここでの説明は省略する。

【0045】以上のように、カメラから、カメラ操作情報や撮影状態情報が得られない場合でも、映像信号を処理することにより、相当する情報を推定することができる。本実施例では、γ補正值、色温度、逆光や過順光の状態などの情報については記載しなかったが、これらの情報も、映像信号を処理することによって獲得することが可能である。上記情報を獲得後、この情報をもとにショットを検出する構成と手法については、第1の実施例と同様であり、説明は省略する。

【0046】（実施例3）次に、本発明のショット検出方法の第3の実施例について説明する。本実施例は、図1のショット変化度合検出部2に関するものである。

【0047】第1の実施例では、図4に示すように、微分フィルタやヒストグラム相関処理によって各情報の変動量を求め、ショット変化度合を検出した。これは、「撮影者が1つのショットを撮り終わって別のショットを撮り始めるまでの期間は、各情報が大きく変化する」、という性質を利用している。しかしながら、入力される情報それぞれについて調べると、情報が大きく変化する現象以外にも、ショットを検出する際に関連するような現象があることがわかる。

【0048】例えば、ある被写体を撮影中に、急に別の被写体にレンズを向けると、フォーカスがずれ、ピンボケの状態が短時間続くことがあるが、この間、高周波成分は、小さくなっている。従って、「高周波成分が小さいとき、別のショットへの移行期間である」というルールが成り立つ。また、同じパンニングでも、ある被写体から別の被写体へ移るためのパンニングは、ショットの移行中と判定してほしいが、ゆっくり移動する被写体を追いかけてパンする場合には、パンニングの期間を1つのショットの区間中として認識したい。このように、同じパンニングでも、ショットの移行中の場合とそうでない場合があるので、第1の実施例のようにパンニング速度の変動量からパンニングの期間を検出できたとしても、ショットの区間を特定できないことがある。しかし、別のショットへ移行するためのパンニングは速くて

短時間でおわるし、被写体をトラッキングする場合には、遅くてほぼ一定速度のパンニングが長時間継続するので、パンニング速度とパンニングの継続時間によって両者を区別することができる。

【0049】このように、ショットが移行する際に各情報がどのような変動パターンを示すかを予め調べておき、この変動パターンに合っているかどうかでショット変化度合を検出することにより、ショットの検出性能を高めることができる。本実施例では、各情報の変動量以外に、各情報の大きさ、各情報が所定の状態を継続する時間、を基にショット変化度合を検出する方法を説明する。

【0050】図9は、高周波成分21とパンニング速度41の2つの情報に関して、ショット変化度合の検出方法を示したものである。図において、微分フィルタ51、絶対値処理部52は、図4の各部と同じであり、67、69は推定部、68は継続時間検出部である。以下に上記構成におけるショット変化度合検出部2の動作を説明する。

【0051】まず、高周波成分21に関する検出方法を説明する。第1の実施例と同様に、微分フィルタ51と絶対値処理部52によって、高周波成分の時間的変動の大きさを求める。推定部67では、高周波成分と絶対値処理部52からの出力を、予め設定した関数で写像し、ショット変化度合として出力する。ここで、写像の関数の具体例として、図10を用いて説明する。図10は、2入力（高周波成分、絶対値処理部52からの出力）、1出力（ショット変化度合）の写像関係を示している。図からわかるように、出力の値域は $[0, \max]$ であり、高周波成分の変動が大きい場合、または高周波成分が小さい場合に大きい値を出力するよう、設定されている。このような写像関数を予め設定することで、2つの入力に対して一意の出力がショット変化度合として求められる。

【0052】次にパンニング速度41に関する検出方法を説明する。この場合も同様にして、微分フィルタ51と絶対値処理部52によって、パンニング速度の時間的変動の大きさを求める。さらに、継続時間検出部68では、微分フィルタ51からの結果を基に、パンニングの継続時間を求める。ここでパンニングの継続時間とは、撮影者がほぼ一定の速度でパンニングしつづけた時間とする。パンニング継続時間の具体的な検出方法としては、まず、パンニング速度の時間的変動量Dが振動許容範囲をはずれた時点から、初めてDが許容範囲内に入る時点をもパンニング開始時点として検出し、パンニング開始時点から初めてDが振動許容範囲をはずれるまでの時間をパンニング継続時間として求める。ただし、振動許容範囲はあらかじめ $[-\text{WIDE}, \text{WIDE}]$ （WIDE：正の所定数）と設定されているものとする。

【0053】推定部69は、パンニング速度と、絶対値

処理部52の出力、パンニング継続時間の3つの値を、予め設定した関数で写像し、ショット変化度合として出力する。ここでの写像の関数は、図10の2入力1出力の関係を3入力1出力に拡張したものであり、「パンニング速度が速いとき」または「パンニング速度の変動が大きいとき」または「パンニング継続時間が短いとき」に、大きい値が出力されるよう設定されている。

【0054】以上のように、ショットが移行する際に各情報がどのような変動パターンを示すかを予め調べておき、この変動パターンに合っているかどうかでショット変化度合を検出することにより、ショットの検出性能を高めることができる。

【0055】なお、上記実施例では、推定部67、69においてショット変化度合を検出する方法として、2入力1出力と3入力1出力の予め設定された写像関数を用いる方法を説明したが、入出力の変換に、例えばファジィ推論や多重線形関数、ニューラルネットなどを用いてもよい。

【0056】また、上記実施例では、入力される情報として高周波成分とパンニング速度の2種類に限定して説明したが、他の情報に対しても同様である。

【0057】また、上記実施例では、各情報の変動量、各情報の大きさ、各情報の所定状態が継続する時間を基に、ショット変化度合の検出を行ったが、入力する情報に応じて、他のパラメータを用いても同様である。

【0058】(実施例4)次に本発明の代表画像記録・表示装置における実施例を説明する。

【0059】本発明の代表画像記録・表示装置は、前述のショット検出方法を利用して、動画像中の代表画像を抽出し、記録・表示するものである。ここで、代表画像とは、ブラウジングや検索などで、動画像を代表する画像として利用するために選ばれる画像であり、撮影者の意図、撮影された画像の状態、被写体の状態などの画像内容が良く表せるように選ばれる。

【0060】図11に本発明の代表画像記録・表示装置の構成図を示す。図において、画像情報出力部1、ショット変化度合検出部2、ショット検出部3、画像処理情報20、カメラ操作情報30、撮影状態情報40は、既に本発明のショット検出方法の実施例で説明したよう

に、図1の構成図と同じである。また、10は代表画像*

$$m(k, i) = \sum_j \text{Info}(k, F(i, j)) / N_i$$

N_i : i番目のショットに属する画像の数

$F(i, j)$: i番目のショットに属する画像の中でj番目の画像

$\text{Info}(k, F(i, j))$: 画像 $F(i, j)$ に関するk番目の入力情報

【0067】と表される。次に、メモリ71は、ショット内代表情報検出部70において検出されたショット毎の代表情報 $M(i)$ を蓄える。代表情報比較部72では、ショット間で代表情報 $M(i)$ の比較を行い、検出された全てのショットの中から「代表画像を出すべきショット

*抽出部、70はショット内代表情報検出部、71はメモリ、72は代表情報比較部、11は映像信号出力部、12は画像記録・表示部である。

【0061】以下、各部の動作を説明する。画像情報出力部1は、画像処理情報20、カメラ操作情報30、撮影状態情報40を出力し、ショット変化度合検出部2において各情報に関するショット変化度合を検出する。ショット検出部3では、複数のショット変化度合を基にして、撮影者が1つのショットとして撮影した画像の区間を特定する。

【0062】一方、代表画像抽出部10は、ショット内代表情報検出部70、メモリ71、代表情報比較部72から構成されており、ショット検出部3で求めたショットの区間を基に、動画像の代表画像を求める。以下、代表画像抽出部10の動作を説明する。

【0063】まず、ショット内代表情報検出部70は、ショット検出部3で検出された各ショットに対して、同一ショットに属する画像の画像処理情報20、カメラ操作情報30、撮影状態情報40を入力し、各情報の平均値を求める。検出された各情報の平均値は「ショットを代表する情報」として出力される。ここで、ショット検出部3で第i番目に検出されたショットに注目し、ショットを代表する情報を数式で表すと、以下のようになる。

【0064】

【数1】

i番目のショットの代表情報 $M(i)$

$$= \{m(1, i), m(2, i), \dots, m(k, i), \dots, m(IN, i)\}$$

【0065】ただし、 IN は、画像処理情報20、カメラ操作情報30、撮影状態情報40の中で入力情報として利用される情報の数とする。例えば、入力情報を、高周波成分、色ヒストグラム、録画開始/終了、ズーム倍率、パンニング速度とする場合、 $IN=5$ であり、 $m(1, i)$ は高周波成分、 $m(2, i)$ は色ヒストグラム、 $m(3, i)$ は録画開始/終了、 $m(4, i)$ はズーム倍率、 $m(5, i)$ はパンニング速度の平均値を示している。従って、k番目 ($0 \leq k \leq IN$) の入力情報に関する代表情報 $m(k, i)$ は、

【0066】

【数2】

ト」を選び、選ばれたショットから代表画像を抽出する。代表情報比較部72の処理手順は以下の通りである。

(1) 最初に検出されたショットを「代表画像を出すべきショット」を選び、ショットを代表する情報 $M(1)$ を

メモリ71内の変数 Mpreに代入して、Mpre=M(1)とする。

(2) 2番目のショットが検出されたとき、ショットを代表する情報 M(2)をMpreと比較し、2つの代表情報間の距離を検出する。ここで代表情報間の距離とは、入力情報毎に1つの次元の軸を割り当てて(IN)次元空間をつくり、この空間に2つの代表情報をプロットしたときの距離である。例えば、a番目のショットとb番目のショットの代表情報の距離Dis(a,b)は、

【0068】

【数3】

$$Dis(a,b) = \left\{ \sum_{k=1}^{IN} (m(k,a) - m(k,b))^2 \right\}^{1/2}$$

【0069】で表される。

(3) Dis(1,2)を所定のしきい値Eと比較し、以下の処理を行う。

・Dis(1,2)>Eのとき、2番目のショットを「代表画像を出すべきショット」として選択し、ショットに属する画像の中から、中央に位置する画像F(2, N2/2)を代表画像として抽出する。また、メモリ71内の変数 Mpreに、代表情報M(2)を代入する。

・Dis(1,2)≤Eのとき、2番目のショットは1番目のショットと画像内容が似ている、と判断される。

(4) 以降で検出されるショットに対しても、(2)、

(3)と同様の処理を行う。すなわち、

(4-1) i番目のショットが検出されたとき、M(i)とMpreと比較し、2つの代表情報間の距離Dis(i,pre)を検出する。

(4-2) Dis(i,pre)をしきい値Eと比較し、

・Dis(i,pre)>Eのとき、i番目のショットを「代表画像を出すべきショット」として選択し、ショットに属する画像の中から、中央に位置する画像F(i, Ni/2)を代表画像として抽出する。メモリ71内の変数Mpreに代表情報M(i)を代入する。

・Dis(i,pre)≤Eのとき、i番目のショットはMpreの情報を持つショットと画像内容が似ている、と判断する。

(4-3) 4-1に戻り、(i+1)番目のショットに対して同様の処理を行う。

【0070】4-1,4-2,4-3の処理は、最後に検出されたショットまで継続され、動画像の代表画像が抽出される。

【0071】一方、代表画像記録・表示部12は、映像信号やその他情報を表示するディスプレイ、または各種情報を記録する記録媒体で構成されている。代表情報比較部72から代表画像に関する情報を入力し、映像信号出力部11からは代表画像に対応する映像信号を入力し、代表画像の表示または記録を行う。

【0072】以上のようにして、ショット間で代表情報を比較し、比較結果を基に選んだショットから、代表画

像を抽出することにより、できるだけ少ない代表画像ですべての画像内容を表示することが可能になる。従って、この代表画像を表示、記録して、動画像のブラウジング、検索などに利用することにより、画像内容が把握しやすくなる。

【0073】

【発明の効果】以上説明したように、本発明のショット検出方法は、カメラ操作情報や画像処理情報や撮影状態情報を入力情報として画像内容の変化する度合を検出することで、撮影者が特定の被写体を撮影しつづけた動画像や特定の画角、撮影条件で撮影しつづけた動画像をショットとして検出できる。

【0074】本発明の方法によれば、画像内容が変化する毎に、動画像が別のショットとして分類されるので、「話の展開が速い」ときは多くのショットが、「ずっと同じような画面が続く」ときには少ないショットが検出され、動画像によって画像内容の変化の仕方が異なる場合にも、画像内容に応じた動画像の単位化が行える。

【0075】また本発明の代表画像表示・記録装置は、上記方法で求めたショットを用いて、代表画像を抽出することにより、少ない代表画像で多くの画像内容を網羅することが可能になり、抽出した代表画像の情報を記録・表示することで、動画像の効率的な検索、早見、編集が行えるようになる。

【図面の簡単な説明】

【図1】本発明のショット検出方法の全体構成を示す図

【図2】動画像のショットとカットの関係を説明するための図

【図3】本発明のショット検出方法の第1の実施例における画像情報出力部の構成を示す図

【図4】本発明のショット検出方法の第1の実施例におけるショット変化度合検出部およびショット検出部の具体的構成を示す図

【図5】本発明のショット検出方法の画像情報出力部の出力の時間変化を示す図

【図6】本発明のショット検出方法の第1の実施例のショット変化度合検出部およびショット検出部に関して、内部ブロックの出力の時間変化を示す図

【図7】本発明のショット検出方法の第1の実施例の画像情報出力部に関して、図3とは別の構成を示す図

【図8】本発明のショット検出方法の第2の実施例の画像情報出力部の構成を示す図

【図9】本発明のショット検出方法の第3の実施例のショット変化度合検出部の構成を示す図

【図10】本発明のショット検出方法の第3の実施例の推定部における2入力1出力の写像関数を示す図

【図11】本発明の代表画像記録・表示装置の実施例の構成を示す図

【符号の説明】

1 画像情報出力部

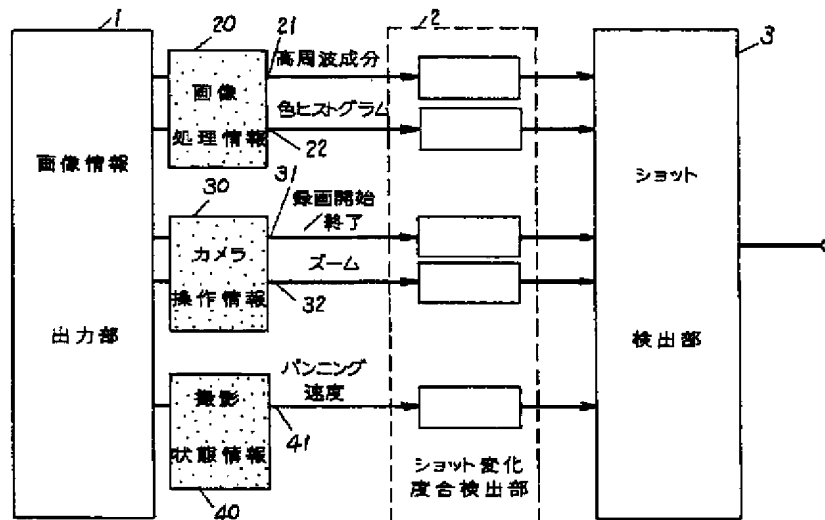
17

18

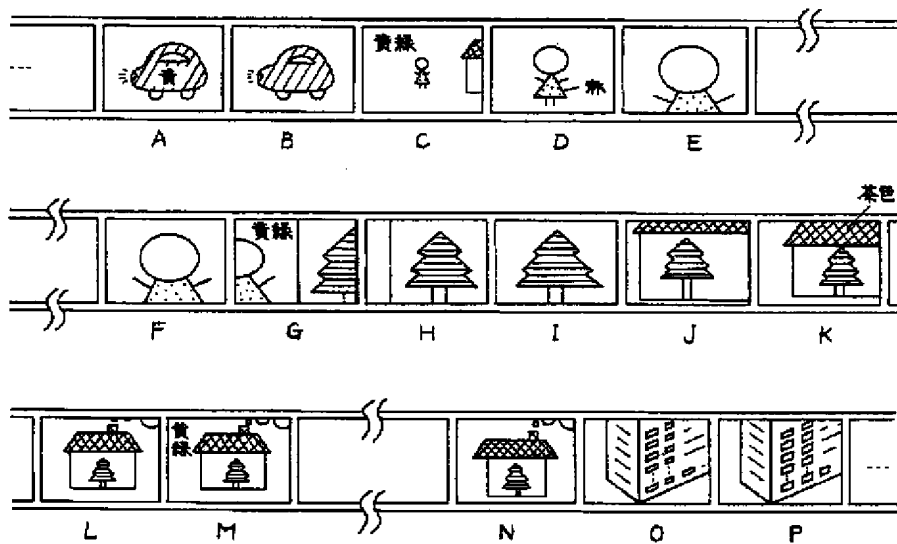
- 2 ショット変化度合検出部
- 3 ショット検出部
- 4 カメラ
- 5 映像信号
- 6 記録媒体
- 7 符号化処理部
- 8 復号化処理部
- 9 画像処理部
- 10 代表画像抽出部
- 11 映像信号出力部
- 12 画像記録表示部
- 20 画像処理情報
- 21 高周波成分
- 22 色ヒストグラム
- 23 動きベクトル
- 24 高周波成分検出部
- 25 色ヒストグラム検出部
- 26 動きベクトル検出部
- 30 カメラ操作情報
- 31 録画開始／終了
- 32 ズーム倍率
- 33 オート／マニュアルのモード
- 40 撮影状態情報
- 41 パンニング速度

- 42 レンズ焦点距離
- 43 絞り開度センサ出力
- 44 フォーカス距離
- 51 微分フィルタ
- 52 絶対値処理部
- 53 ローパスフィルタ
- 54 ゲイン調整部
- 56 最大値検出部
- 57 状態判定部
- 10 58 ショット区間出力部
- 59 カウンタ
- 60 フレーム間差分値検出部
- 61 メモリ
- 62 変化量検出部
- 63 カットチェンジ検出部
- 64 カメラワーク検出部
- 65 動きベクトル検出部
- 66 カメラワークパラメータ推定部
- 67、69 推定部
- 20 68 継続時間検出部
- 70 ショット内代表情報検出部
- 71 メモリ
- 72 代表情報比較部

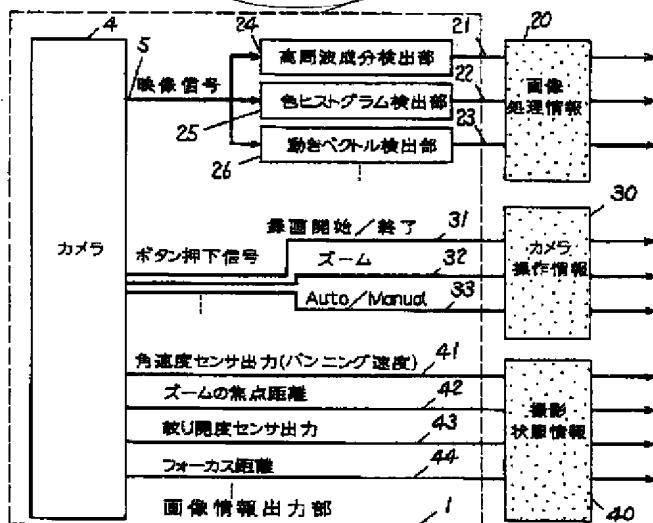
【図1】



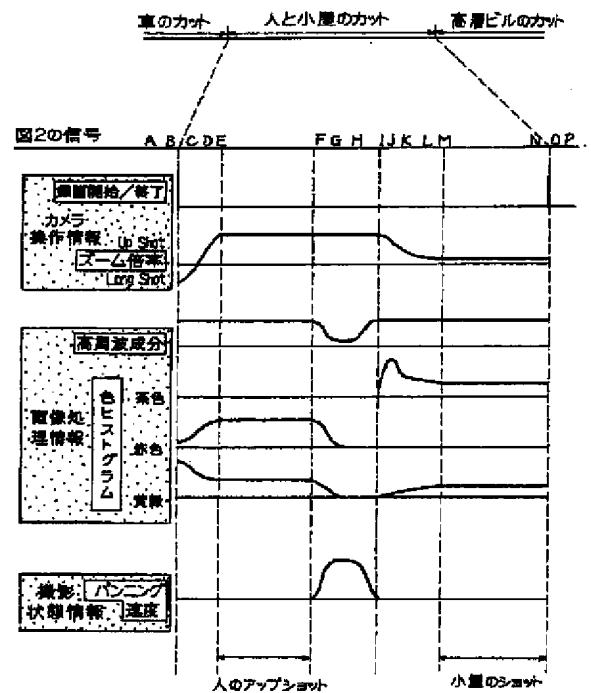
【図2】



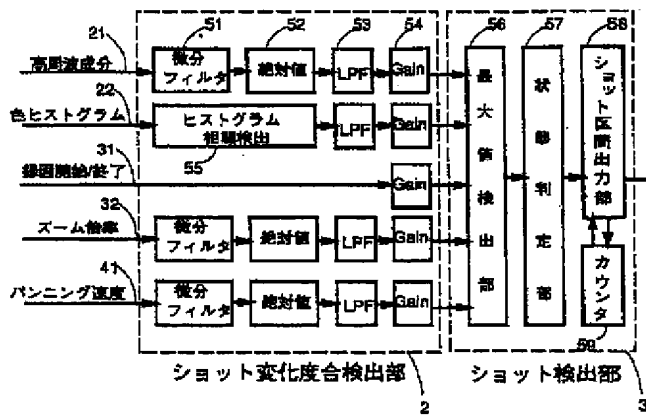
【図3】



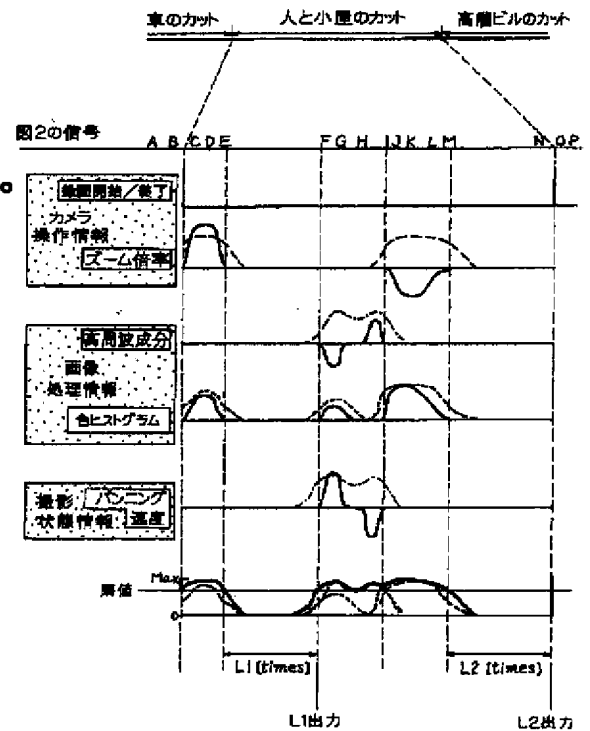
【図5】



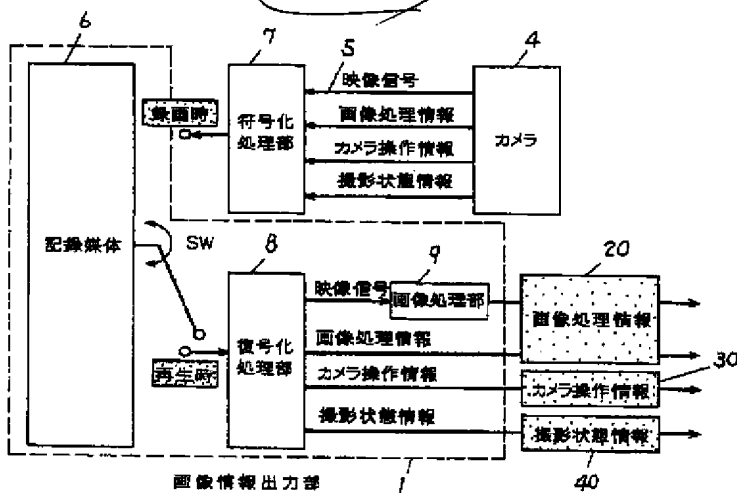
【図4】



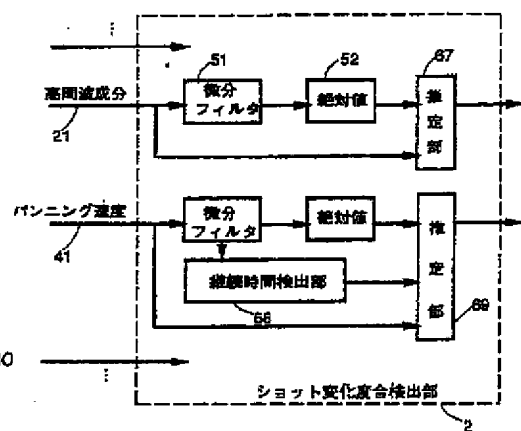
【図6】



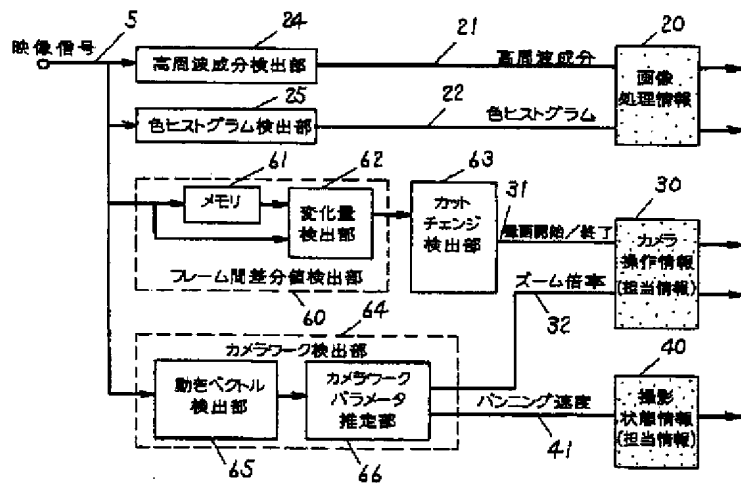
【図7】



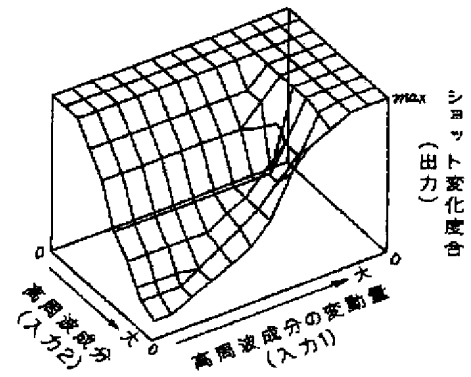
【図9】



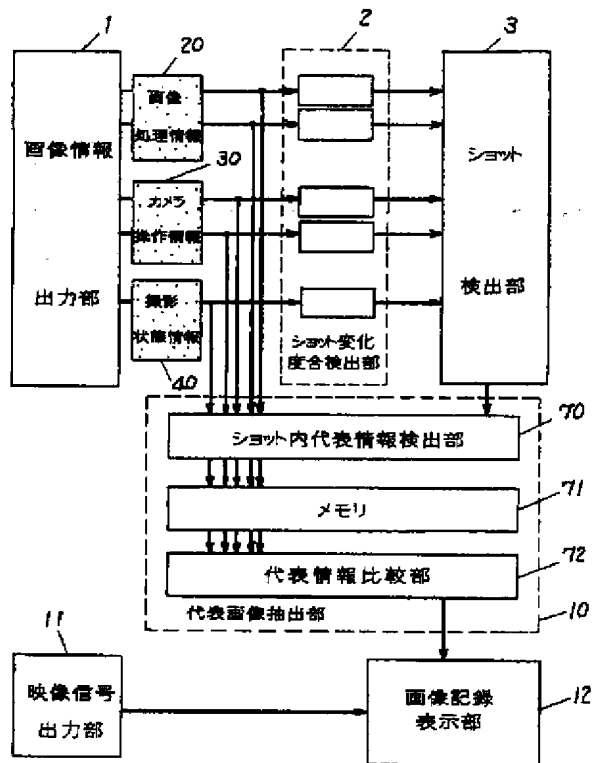
【図8】



【図10】



【図11】



フロントページの続き

(51) Int. Cl.⁶

識別記号

庁内整理番号

F I

技術表示箇所

H 0 4 N 5/782

K